



OPEN ACCESS

EDITED BY

Andrew Cohen,
Drexel University, United States

REVIEWED BY

Paula Sampaio,
Universidade do Porto, Portugal
Felix Yuran Zhou,
University of Texas Southwestern Medical
Center, United States

*CORRESPONDENCE

Leila B. Mostaço-Guidolin,
✉ leila@sce.carleton.ca

[†]These authors have contributed equally to
this work and share first authorship

RECEIVED 05 December 2024

ACCEPTED 19 March 2025

PUBLISHED 17 April 2025

CITATION

Kunchur NN, Poole JJA, Levine J, Hackett T-L,
Thornhill R and Mostaço-Guidolin LB (2025)
Classification of collagen remodeling in
asthma using second-harmonic generation
imaging, supervised machine learning and
texture-based analysis.
Front. Bioinform. 5:1539936.
doi: 10.3389/fbinf.2025.1539936

COPYRIGHT

© 2025 Kunchur, Poole, Levine, Hackett,
Thornhill and Mostaço-Guidolin. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with
these terms.

Classification of collagen remodeling in asthma using second-harmonic generation imaging, supervised machine learning and texture-based analysis

Natasha N. Kunchur^{1†}, Joshua J. A. Poole^{1†}, Jesse Levine¹,
Tillie-Louise Hackett², Rebecca Thornhill^{3,4} and
Leila B. Mostaço-Guidolin^{1*}

¹Department of Systems and Computer Engineering at Carleton University, Ottawa, ON, Canada, ²Anesthesiology, Pharmacology and Therapeutics Department at the University of British Columbia, Medical Sciences, Vancouver, BC, Canada, ³Department of Radiology, Radiation Oncology, and Medical Physics at the University of Ottawa, Ottawa, ON, Canada, ⁴Department of Medical Imaging at the Ottawa Hospital, Ottawa, ON, Canada

Airway remodeling is present in all stages of asthma severity and has been linked to reduced lung function, airway hyperresponsiveness and increased deposition of fibrillar collagens. Traditional histological staining methods used to visualize the fibrotic response are poorly suited to capture the morphological traits of extracellular matrix (ECM) proteins in their native state, hindering our understanding of disease pathology. Conversely, second harmonic generation (SHG), provides label-free, high-resolution visualization of fibrillar collagen; a primary ECM protein contributing to the loss of asthmatic lung elasticity. From a cohort of 13 human lung donors, SHG-imaged collagen belonging to non-asthmatic (control) and asthmatic donors was evaluated through a custom textural classification pipeline. Integrated with supervised machine learning, the pipeline enables the precise quantification and characterization of collagen, delineating amongst control and remodeled airways. Collagen distribution is quantified and characterized using 80 textural features belonging to the Gray Level Cooccurrence Matrix (GLCM), Gray Level Size Zone Matrix (GLSZM), Gray Level Run Length Matrix (GLRLM), Gray Level Dependence Matrix (GLDM) and Neighboring Gray Tone Difference Matrix (NGTDM). To denote an accurate subset of features reflective of fibrillar collagen formation; filter, wrapper, embedded and novel statistical methods were applied as feature refinement. Textural feature subsets of high predictor importance trained a support vector machine model, achieving an AUC-ROC of $94\% \pm 0.0001$ in the classification of remodeled airway collagen vs. control lung tissue. Combined with detailed texture analysis and supervised ML, we demonstrate that morphological variation

amongst remodeled SHG-imaged collagen in lung tissue can be successfully characterized.

KEYWORDS

airway remodeling, asthma, collagen, machine learning, second harmonic generation, texture analysis

1 Introduction

ASTHMA is characterized by chronic airway inflammation (the target of all current asthma therapeutics) and airway remodeling, which involves the epithelium, basement membrane, lamina propria, smooth muscle and vascular structures (Holgate, 2008). Airway remodeling was first denoted to be present in fatal asthmatics by Hubert and Kossler in 1922 and since then, has been documented in all stages of asthma severity (Huber, 1922). Airway remodeling has also been linked to reduced lung function, airway hyperresponsiveness and greater use of asthma medications (Bergeron et al., 2010; Boulet et al., 1997; Elias et al., 1999; James et al., 2002; Zeiger et al., 1999). Increased deposition of fibrillar collagens-I and III has been shown within the extracellular matrix (ECM) remodeled airways of asthmatic patients (Dunnill, 1960).

There is a natural abundance of fibrillar collagen within the ECM of the lung to retain tensile strength and support suitable architecture. However, alterations in the concentration and distribution of collagen fibres can lead to loss of anatomical structure, compromised cell function, and tissue fibrosis (Buehler, 2006; Engler et al., 2009; Hoshino et al., 2010; Liu L. et al., 2021).

To understand the changes in ECM at a microcellular level, tissue analysis is conducted ex-vivo. Historically, histological staining has been the predominant approach for studying the complexity of airway remodeling and for visualizing alterations in cell behaviour and tissue structure (Dunnill, 1960; Mostaço-Guidolin et al., 2021). The abundant proliferation of collagen in asthmatic airway remodeling was first visualized using Verhoeff-van Gieson and Masson trichrome staining (Carroll et al., 2000; Dunnill, 1960; Hogg, 1997). While histological staining remains the gold standard for assessing tissue alterations, the process of staining and sample fixation often alters the tissue structure through protein denaturation and cross-linking (Alturkistani et al., 2015; Javaeed et al., 2021). Furthermore, the representation of the collagen fibers are mixed with other extracellular components, making it difficult to resolve exact collagen proteins within a sample when applying H&E staining (Keikhosravi et al., 2020). These traditional staining methods lack specificity for detailing ECM proteins' biochemical and structural information, particularly the fibrillar collagen deposited during airway remodeling.

A non-centrosymmetric structure in crystallography refers to a crystal structure that lacks a center of symmetry. In such structures, there is no point within the crystal through which every part of the structure has an identical part located symmetrically opposite. This lack of inversion symmetry often leads to unique physical properties, such as piezoelectricity, second-harmonic generation, and other nonlinear optical effects.

Fibrillar collagen lacks a centre of symmetry, demonstrating a non-centrosymmetric property. This lack of inversion symmetry leads to its unique physical properties, making it an ideal candidate for nonlinear optical effects such as second harmonic generation (SHG) imaging (Bancelin et al., 2014; Chen et al., 2012; Cox et al., 2003; Han et al., 2005; Poole Mostaço-Guidolin et al., 2021). Since the first demonstration of the SHG application and its capabilities in imaging biomolecules, it has been widely used to monitor fibrillar collagen and its role in remodeling and wound repair (Deka et al., 2012; Israel et al., 2017; Sun et al., 2008; Tilbury et al., 2014). SHG allows for label-free tissue structure assessment without requiring exogenous labelling and/or extensive sample preparation. Visualizing fibrillar collagen through SHG imaging has provided insight into the morphological alterations in asthmatic lung tissue (Mostaço-Guidolin et al., 2019).

With the promise of non-invasive imaging, SHG is capable of visualizing freshly obtained tissues in their natural physiological state. With this, the structural information of the samples can be accessed immediately. Additionally, when combining fixation methods such as paraffin embedding with SHG microscopy, one can achieve prolonged storage of tissue, increasing SHG microscopy versatility to image fresh or preserved tissues. Fixation allows for maintaining and highlighting the integrity of tissue morphology. Furthermore, paraffin-embedding supports the serial sectioning of tissues, allowing for detailed cross-sectional analysis of tissues, inclusive of airways. This routinely adopted method can guide digital reconstruction of 3D data and visualize the intricate anatomy of the lungs (the bronchioles, alveoli and alveolar capillary network) (Baschong et al., 2001; Grothausmann et al., 2017; Monaghan et al., 2016). Furthermore, paraffin-fixed tissues allow for successful reinvestigation of archived human tissue. Authors Bredfeldt et al., follow a similar approach, where collagenous fibres of cancerous breast-tissue are evaluated using SHG upon having been paraffin-embedded and H&E stained (Bredfeldt et al., 2014). Different groups have explored and tested the feasibility of using paraffin-embedded tissues for both stained-based microscopy and label-free imaging (Martin et al., 2013; Sabo et al., 2022; Vijayaraghavan et al., 2014). Findings support the use of paraffin-guided fixation agents, countering the common perception that these tissue additives introduce harmful autofluorescence when imaging. Sabo et al., successfully retrieve specialized high-content data from paraffinized tissue sections when measuring collagen distribution in interstitial fibrosis with SHG imaging (Sabo et al., 2022).

However, to fully capitalize on the intricate level of detail offered by SHG imaging, effective pattern recognition tools are required to appreciate the specific features associated with tissue remodeling. Gray-level textures represent an advanced statistical approach for quantifying the spatial distribution of gray-level pixel intensities within a localized area (Liu Y. et al., 2021; Song et al., 2022). With the organization of pixel intensities

into a tabular format, various features detailing the structural properties of the SHG-imaged tissue can be derived through the quantitative distribution of pixels. We explore the application of five primary texture groups: Gray Level Cooccurrence Matrix (GLCM), Gray Level Size Zone Matrix (GLSZM), Gray Level Run Length Matrix (GLRLM), Gray Level Dependence Matrix (GLDM), and Neighboring Gray Tone Difference Matrix (NGTDM) (Amadasun and King, 1989; Galloway, 1975; Haralick et al., 1973; Sun and Wee, 1983; Thibault et al., 2014). Through the examination of pixel arrangement, gray-level textures can help infer the underlying organization and structure of collagen fibrils (Cicchi et al., 2013; Hu et al., 2012; Wu et al., 2016). The extracted textural features are used to train a support vector machine capable of differentiating between control and remodeled asthmatic airways. In automating the classification of SHG images based on remodeling status, we present a quantitative approach for monitoring changes using gray-level textures. We demonstrate SVM's robust capabilities as a powerful classification tool adept at discerning morphology types within collagen fibrils, while successfully accounting for minute differences easily missed by the human eye. This method offers scalability, allowing for its extension across various contexts and pathologies.

With the interactions between the microenvironment and its cellular components slowly becoming understood as a driving force in several chronic diseases, our work presents a methodology capable of visualizing these imposed structural changes during the onset of asthma. Our developed pipeline proposes a proof-of-concept that explores the non-invasive characterization of tissue remodeling and paves the way to assess detail regarding ECM structure alterations which are not possible to track using common staining strategies. An increased understanding of the ECM and its contribution in chronic lung diseases can support the testing and development of novel targeted therapies. While leveraging SHG microscopy and robust image analysis methods, we propose a novel image-based, reproducible quantitative methodology pipeline, capable of exploring diseased microenvironments and tracking changes associated to tissue structure alterations due to diseases and/or therapeutic interventions.

2 Materials and methods

2.1 Human lung tissue preparation and dataset

A total of 12 lungs deemed unsuitable for transplantation were obtained with informed consent from next-of-kin of donors with asthma ($n = 6$) and donor controls ($n = 6$) through the International Institute for the Advancement of Medicine (IIAM, Edison, NJ; www.iiam.org) and biobanked within the James Hogg Lung Registry (Ethics protocol number H00-50110). This study was approved by the Providence Healthcare Research Ethics Board (H13-02173). As previously described, the lungs were air-inflated to 10 cmH₂O, and frozen over liquid nitrogen vapour. Lungs were then cut into 2 cm thick slices (transaxial plane) (Hogg et al., 2004; Mostaço-Guidolin et al., 2019) and each slice was uniformly randomly sampled by using a line grid (225 mm²), superimposed onto the photographs of each lung slice. Sampled cylindrical tissue

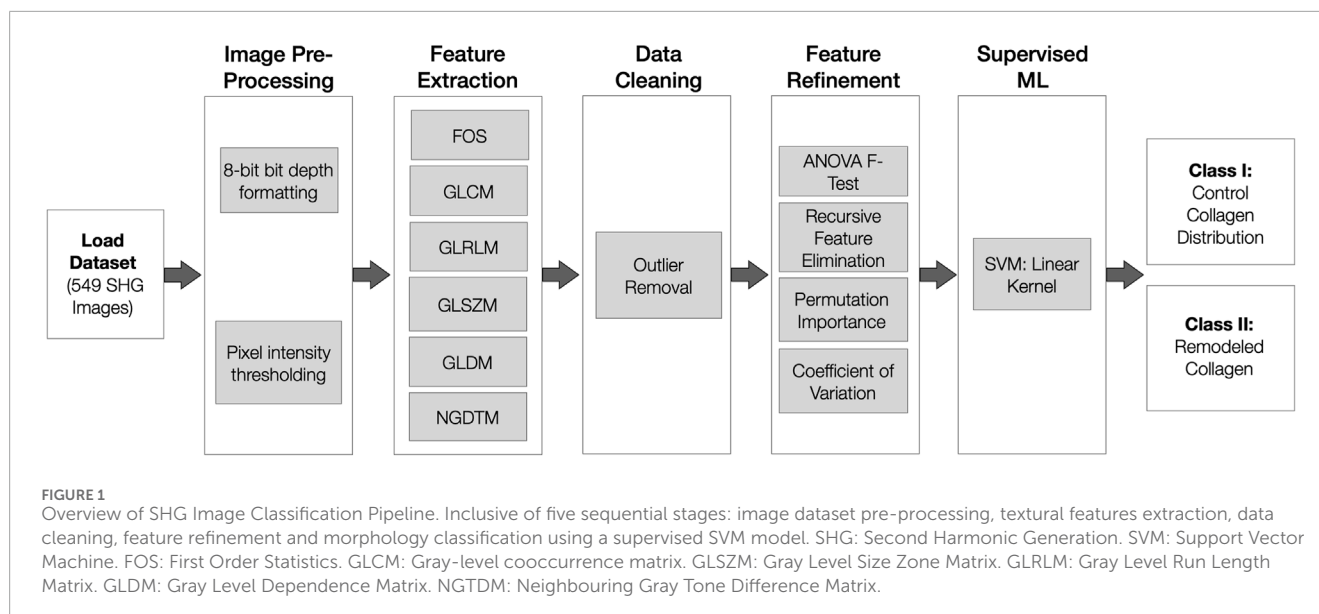
cores (15 × 20 mm) were formalin-fixed and paraffin-embedded. Sequential 5 µm tissue sections containing airways were used for SHG imaging, with sectioned counterparts stained with Verhoeff-van Gieson stain and digitally scanned (ScanScope XT slide scanner; Aperio Technologies) for reference. Through sequential slicing, the orientation of the tissue sections remained consistent throughout imaging. Of the total dataset of 549 images; 296 are representative of asthmatic donors and 253 representative of control. The exact image collection amongst control and asthma is as follows: control (31, 37, 37, 43, 47, 58) and asthma (43, 55, 38, 39, 56, 65).

2.2 Second harmonic generation (SHG) imaging

Each unstained tissue section was imaged using a multimodal Nonlinear Optical microscope (NLOM), housed at the National Research Council in Winnipeg, Canada. The custom-built system allowed for the acquisition of correlative two-photon excitation fluorescence (TPEF), Coherent Anti-Stokes Raman Scattering (CARS) and SHG imaging, as previously described (Ko et al., 2009; Ko et al., 2010). A Ti:Sapphire oscillator (Tsunami, Spectra-Physics) operating at 800 nm with a 100 fs pulse duration and 20 nm bandwidth was used. The backward SHG signal was collected and transmitted to a non-descanned PMT detector (H9656 series, Hamamatsu, Bridgewater, NJ, USA) mounted on the microscope assembly (400 ± 30 nm), using a 20x, 0.75NA infinity corrected air objective lens (Olympus Canada, Markham, ON, Canada). The system generated images of approximately 450 × 450 µm with a translational resolution of 0.2 µm. A motor controller (MP-285, Sutter Instrument, Novato, CA, USA) provided the motorized translational movement of the stage (Sutter Instrument, Novato, CA, USA). ScanImage (ver.3.5) software was used for image acquisition and laser scanning control (Pologruto et al., 2003). The typical pixel dwell time for an average of four scans for a single frame collection was 21 µs.

2.3 SHG image classification pipeline overview

A total of 549 labelled SHG images of airways were obtained and used in this study. The differences in signal distribution patterns associated with collagen fibre deposition were characterized using the image classification pipeline shown in Figure 1. The developed pipeline is comprised of five chronological steps: image preprocessing, feature extraction, data cleaning, feature refinement and classification. To support adequate textural analysis, image preprocessing first ensures the SHG images are of 8-bit bit-depth and TIFF format, preserving the quality of the graphics. To represent a standardized intensity of SHG fibrillar collagen signal amongst all included images, the dataset had to meet a minimum pixel intensity threshold. For each of the images, 5 gray-level texture matrices are obtained in parallel, extracting relevant features capable of capturing the unique pixel arrangement and characteristics of collagen distribution. The extracted features are sequentially cleaned, replacing outlier data with median values. A total of 80 features were extracted per image and refined using i) ANOVA F-test, ii) recursive feature elimination, iii) permutation importance and iv) coefficient of variation strategies. Lastly, the refined features



were supplied into two supervised machine learning models in parallel; testing which framework achieved the highest classification accuracy when discriminating amongst control and remodelled airway tissue.

2.4 Image preprocessing

In this study, SHG images originally exported as 16-bit TIFF files were converted to 8-bit in ImageJ (Schneider et al., 2012), for texture analysis, reducing both computational demands and memory usage while still preserving the core textural information (Bankfield, 2024). Although this down conversion restricts the dynamic range, it was deemed acceptable given our focus on relative texture features rather than absolute intensities. A predefined intensity threshold ($I < 5$) was then applied to exclude low-signal images that lacked sufficient grey-level information for reliable textural analysis. To address outlier values, we replaced anomalous intensities with the median value of the dataset; although alternative methods exist, median replacement was selected for its simplicity and reproducibility when processing large image sets.

2.5 Feature extraction and data cleaning

To evaluate the textural features of collagen distribution, a total of 80 gray-level texture features were extracted from each image, including four first-order statistic (FOS) features, 24 Gray-level cooccurrence matrix (GLCM) features, 16 Gray Level Size Zone Matrix (GLSZM) features, 16 Gray Level Run Length Matrix (GLRLM) features, 15 Gray Level Dependence Matrix (GLDM) features, and 5 Neighboring Gray Tone Difference Matrix (NGDTM) features. The matrix operations and the respective features extracted for each family of texture features are defined in [Supplementary Table S1](#). Gray-level texture analysis was performed computationally in two steps: i) base construction of each of the gray-level texture matrices, and ii) calculation of features from

each of the matrices. For the GLCM and GLRLM, all features extracted were averaged from four directional matrices at 0° , 45° , 90° and 135° . To address outlier cases, the interquartile range method (IQR) was applied due to its simplicity and low susceptibility to noise (Barbato et al., 2011). While effective for distributions with high skew, it lacks the adaptability to individual cases that other outlier rejection methods might possess. Given the inability to assume that each feature's distribution is free of high skew, the IQR method emerged as a suitable choice for outlier rejection. Rather than outright removal from the dataset, rejected outliers were substituted with the median value. The median is a robust measure of central tendency and does not shift significantly with outliers such as the mean. Bias is minimal when using median values as the replacement value is independent of skew, preserving the original data structure. As our data is non-normally distributed, using the median ensures that the imputed values are representative of the data. Furthermore, the median value accommodates both small and large datasets (Maronna et al., 2006). This approach is aimed at preserving individual data points and not further limit the size of our dataset. With a comprehensive initial list of 80 texture features extracted from the SHG dataset, we next focus on dimensionality reduction and feature refinement to avoid overfitting and ensure our models capture the most predictive information, as detailed in Section F. This preserves data quality, integrity and further support the reproducibility of our methodology. All methods applied to feature extraction and data cleaning were performed using custom “in-house” Python scripts.

2.6 Feature refinement

With 80 features, dimensionality reduction and feature refinement are required to avoid overfitting and to allow adequate generalizability of our collagen classification models. To ensure the high predictive power of the supplied features, four feature refinement methods were investigated: ANOVA F-test (filter

method), recursive feature elimination (wrapper method), permutation method (embedded method) and the novel coefficient of variation (statistical) method. The ANOVA F-Test identifies features with the highest ratio of variance between different groups and within the same groups. Distributions that exhibit minimal overlap amongst features are selected, to provide the model with distinctive features that can facilitate easier discrimination between two groups of data (Dissanayake Md Johar, 2021; Remeseiro and Bolon-canedo, 2019). The top ten features ranked using ANOVA F-test scores were selected. Recursive Feature Elimination (RFE) deploys all features initially into a decision tree model, progressively removing one feature at a time while re-fitting the defined decision tree model (Guyon et al., 2002). The iterative process continues until a drop in model accuracy is observed, indicating that the optimal number of features has been reached. Alternatively, permutation importance (PI) determines the optimal feature subset by establishing a baseline score through the training of a random forest classifier. During training, the model is continuously fitted to accommodate a feature which has been “permuted/rearranged”. The random forest classifier is employed to make predictions using this permuted validation data and re-evaluated (Altmann et al., 2010). The PI score for a feature is computed as the difference between the baseline score and the score achieved when the model predicts using the permuted feature. This permutation process is repeated for each feature over 100 iterations to determine which features, on average, cause the most significant changes in accuracy (Altmann et al., 2010). Lastly, a novel Coefficient of Variation (CV) method was applied to capture interclass feature variation. The CV reflects the standard deviation from the mean, where a higher CV signifies a greater dispersion of the mean. As demonstrated in a previous application (Kunchur et al., 2022), a CV threshold of 8% was applied. The detailed methodology of the CV method can be found elsewhere. The subsets of features extracted from each refinement technique were independently used to train a support vector machine model. ANOVA F-test, PI and RFE feature refinement were implemented using Scikit Learn packages in Python (scikit-learn, developers BSD L, 2007a; scikit-learn, developers BSD L, 2007b; scikit-learn developers, BSD License, 2007; scikit-learn, developers BSD L, 2007c).

2.7 Supervised machine learning

We trained support vector machine models to delineate SHG images obtained from non-asthmatic controls from those extracted from remodeled (asthmatic) airways. Through the optimization of a hyperplane providing maximal class separation, SVMs have proven to be powerful binary classification models. The default parameters of the sklearn.svm model with a linear kernel were used (scikit-learn, developers BSD L, 2007c). Both linear and RBF kernels were compared, however, linear presented stronger classification metrics. With a balanced dataset, the class weight function was set to a default parameter of one. Linear discriminant analysis was used to further reduce the dimensionality of the feature set for each of the feature refinement methods prior to SVM training. Feature refinement methods and LDA were used in unison to reduce redundant/overlapping features amongst the five textural matrices employed. Feature refinement methods were leveraged with the

intention to remove similar features; with LDA then identifying the best discriminative space. When used in combination, higher classification metrics (the ones presented) were observed. The accuracy of each model was assessed through Repeated Stratified K-Fold validation, employing ten splits and conducting three runs per split. Accuracy scores were computed for each fold, and the mean of these scores served as the ultimate accuracy metric for the model. Each model underwent training on a randomized split, utilizing 66% for training data and 33% for test data, with a fixed random state of one. The split was conducted at the image level, through the randomization of the samples amongst donors. Precision, recall, and F1-score were documented, alongside the accuracy from Repeated Stratified K-Fold validation. This classification process was repeated for each set of feature subsets determined by the previously mentioned feature refinement methods, inclusive of a final aggregation of all selected features. All supervised ML models, dimensionality reduction, cross-validation and performance scoring were implemented using Python3 SciKit Learn ML packages.

3 Experimental results

Representative SHG images depicting the collagen organization in non-asthmatic control and an asthmatic airway are depicted in Figure 2.

3.1 Performance using all 80 extracted features

Though it is not advisable to supply simple supervised classification models with many more features than positive (asthmatic) cases available for classification, our initial objective was to investigate the functionality of using all 80 extracted gray-level textures. An overall k-fold cross-validation accuracy of $88\% \pm 1\%$ in the classification of asthmatic airways vs. control airways was observed with a macro-averaged ROC-AUC of 0.98 ± 0.0 , for $k = 10$, as summarized in Supplementary Figure S1. The model exhibited a higher precision of 0.88 ± 0.0 reflected in the classification of asthmatic airways in comparison to control airways.

3.2 Performance using refined Feature sets

Four feature refinement methods were tested and four SVM models were independently trained on the feature subsets identified by each refinement method. The selected features of each refinement method are shown in Table 1.

The performance achieved by each trained model on the four feature subsets are shown in Figure 3 and Table 2. Refined features selected by ANOVA F-Test (Filter Method) and Recursive Feature Elimination (Wrapper method) demonstrated high predictive importance, presenting optimal classification metrics of $86\% \pm 1\%$ and $87\% \pm 2\%$ accuracies respectively. With a selection of only seven features in total, RFE was capable of delineating between remodeled and control airways. SVM trained on RFE features boasted a k-Fold Cross Validation accuracy of $87\% \pm 2\%$ with a Macro-Averaged ROC-AUC of 0.94 ± 0.0 . The ANOVA

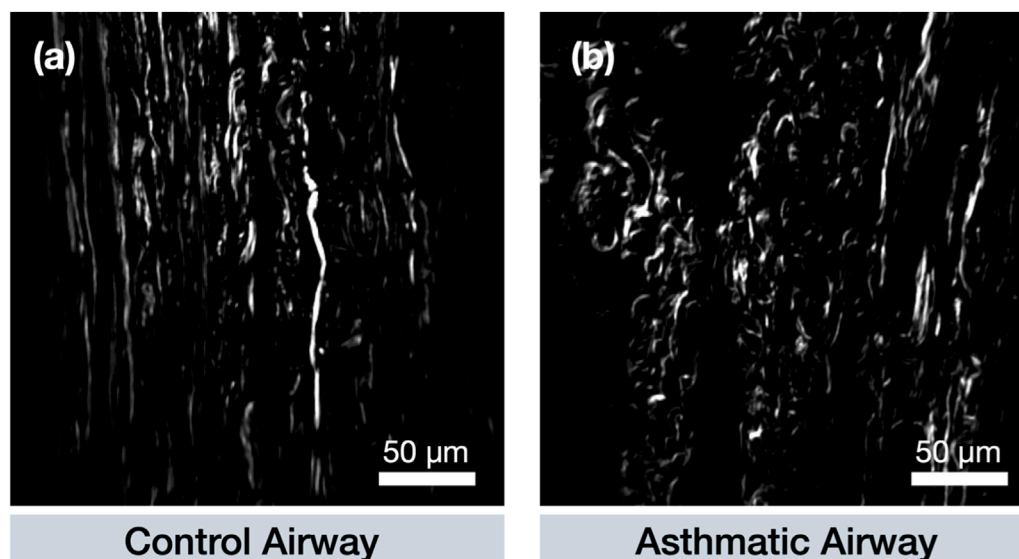


FIGURE 2

Representative SHG images of fibrillar collagen deposition in airways from (a) control and (a) asthmatic donors. A more uniform collagen fibre arrangement is evident in control airways, while a more fragmented structure is typically present in asthmatic cases. (a). Performance using all 80 extracted features.

F-test achieved a similar performance, with k-Fold Cross Validation accuracy of $86\% \pm 1\%$ and Macro-Averaged ROC-AUC of 0.94 ± 0.0 . In comparison, models trained on features selected by PI and CV achieved lower accuracies, with respective K-Fold Cross Validation accuracies of $80\% \pm 1\%$ and $82\% \pm 1\%$ and ROC-AUCs of 0.86 ± 0.0 and 0.88 ± 0.0 . Confusion matrices for each classifier are found in [Supplementary Figure S2](#).

3.3 Performance using Pooled set of 33 features

To improve classification accuracy and mitigate the risk of overfitting, the final subset of tested features comprised a pooling of all features identified using the four feature refinement methods. Redundant features common to all methods were eliminated, establishing a subset of 33 unique gray-level textural features (listed in [Supplementary Tables S4, 5](#)). High classification performance metrics of asthmatic and control collagen morphologies were observed, with a k-fold cross-validation accuracy of $86\% \pm 1\%$ and Macro-Averaged ROC-AUC of 0.97 ± 0.0 . The confusion matrix for pooled features is shown [Supplementary Figure S3](#). Concerning the F1 scores in [Figure 4](#), the model exhibited an F1-score of 0.83 for NA (control) indicating the model's difficulty in accurately predicting the positive instances of NA (control).

4 Discussion

Insight into the structural alterations in ECM is essential to understanding disease pathology. In this study, we demonstrate the utility of supervised machine learning coupled with texture-based

analysis in elucidating ECM remodeling in asthma. Specifically, we highlight that the deposition and morphology of fibrillar collagen in control and remodeled airways vastly differ. We apply first- and second-order statistical texture features to quantitatively assess the variance in collagen morphology between non-asthmatic control and asthmatic or remodeled airway states. By identifying distinct traits specific to control and remodeled collagen, we can objectively track ECM modifications in a diseased environment, specifically in asthma. Given that the interactions between the microenvironment and its cellular counterparts are slowly becoming understood as a driving force between tissue and chronic disease, there is a need to establish quantitative methodologies capable of tracking said morphological changes.

Gray-level texture analysis presents an effective approach to monitoring tissue ECM alterations, especially in the structure and morphology of the collagen network. In a broader context, the five textural matrices have been commonly applied independently to track textural features denoting changes in disease states of varying imaging modalities such as MRI, CT images, ultrasound, x-ray microscopic cellular and histopathological images ([Alvarenga et al., 2010](#); [Chabat et al., 2003](#); [Gibbs and Turnbull, 2003](#); [Herlidou et al., 2004](#); [Moura et al., 2022](#)). Success in the evaluation of collagen deposition was demonstrated using SHG and GLCM in tracking the aging of tissues and the onset of pancreatic cancer ([Hu et al., 2012](#)). To provide an exhaustive overview of applicable features, features derived from five gray-level texture families (GLCM, GLRLM, GLDM, GLSZM, NGTDM) were evaluated. Though the operation of some of the matrices is deemed unconventional for tracking biological systems, the mathematical approach for feature characterization can easily be extrapolated to evaluate any convention, providing novelty to the analysis conducted. A combination of said textural matrices has been applied in examples

TABLE 1 Textural features included after the application of each feature refinement method.

Matrix	GLCM	NGTDM	GLSZM	GLRLM	GLDLM
ANOVA F Test - Filter Method (10 features)					
Feature	• Correlation • Maximal Correlation Coefficient	• Busyness	• Large Area Emphasis • Large Area Low Gray Level Zone Emphasis • Size Zone Variance	• Short Run Low Gray Emphasis • Gray Level Non-Uniformity Normalized • Gray Level Non-Uniformity • Run Entropy	• (no selected features)
Permutation Importance (PI) – Embedded Method (9 features)					
Feature	• Joint Entropy • Difference Entropy • Inverse Difference Normalized	• (no features selected)	• Small Area Emphasis • Size Zone Non-Uniformity Normalized	• Run Variance • Long Run Low Gray Emphasis • High Gray Level Run Emphasis	• Low Gray Level Emphasis
Recursive Feature Elimination (RFE) - Wrapper Method (7 features)					
Feature	• Correlation	• (no features selected)	• Large Area Emphasis • Large Area Low Gray Level Zone Emphasis • Size Zone Percentage	• Short Run Low Gray Emphasis • Gray Level Non-Uniformity Normalized • Run Entropy	• (no features selected)
Coefficient of Variation Method (CoV) - (17 features selected)					
Feature	• Autocorrelation • Cluster Prominence • Cluster Shade • Cluster Tendency	• Coarseness • Busyness	• Large Area Emphasis • High Gray Level Zone Emphasis • Small Area High Gray Level Zone Emphasis • Gray Level Variance • Size Zone Variance	• Short Run High Gray Emphasis • Gray Level Variance • Run Variance	• High Gray Level Emphasis • Small Dependence High Gray Level Emphasis • Gray Level Variance

GLCM: Gray-level co-occurrence matrix. GLSZM: Gray Level Size Zone Matrix. GLRLM: Gray Level Run Length Matrix. GLDLM: Gray Level Dependence Matrix. NGTDM: neighbouring gray tone difference matrix.

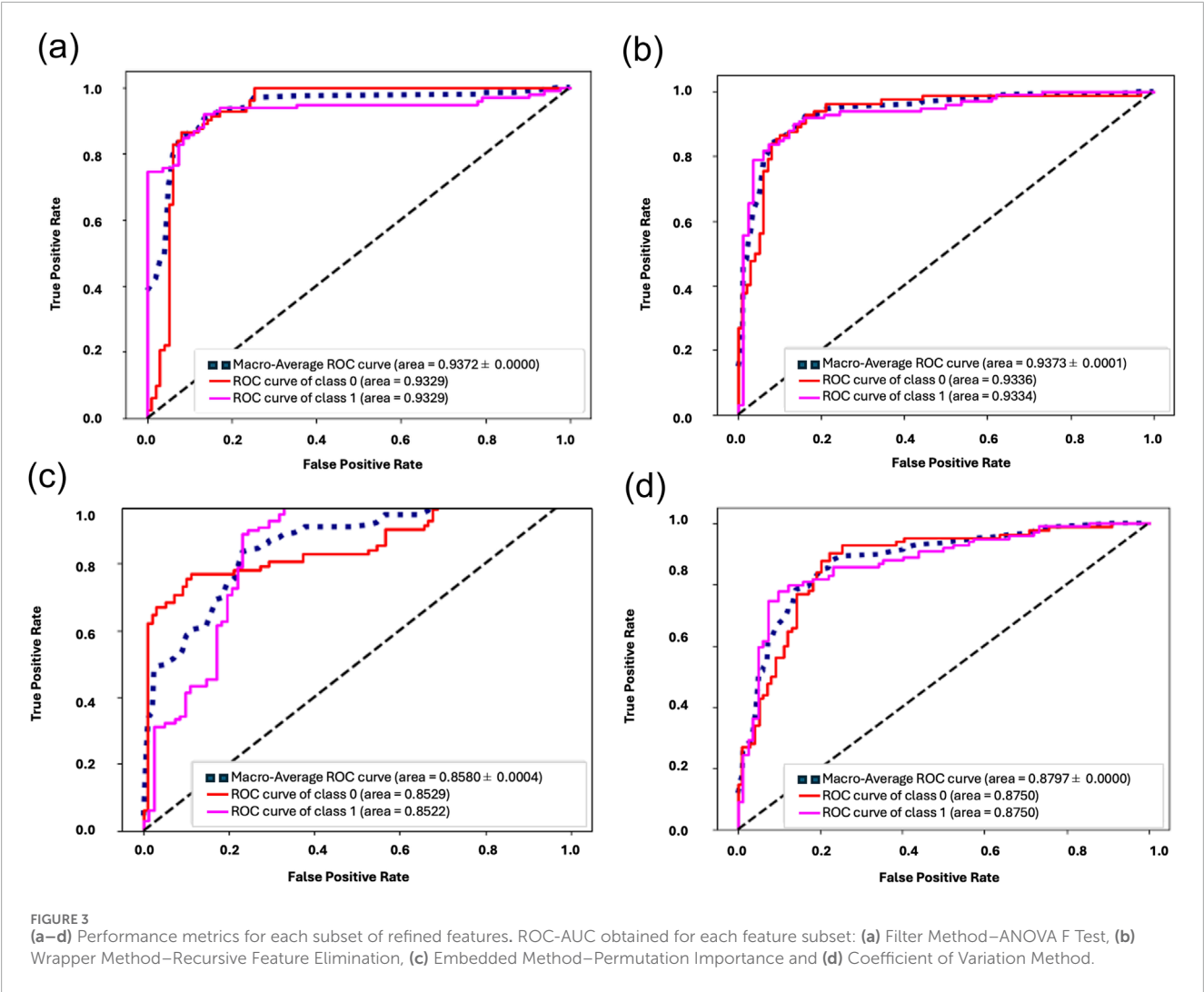
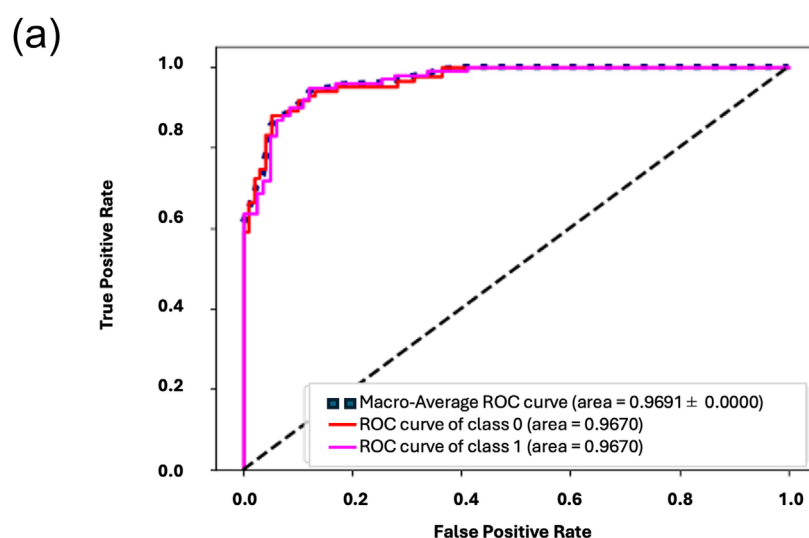


TABLE 2 Summary of performance metrics of SVM trained on each of the four distinct feature subsets.

Method	Status	Precision	Recall	F1-score	Best Accuracy
Filter - ANOVA F	Control	0.84	0.83	0.84	86%
	Asthmatic	0.88	0.88	0.88	
Wrapper – RFE	Control	0.84	0.83	0.84	87%
	Asthmatic	0.88	0.88	0.88	
Embedded – PI	Control	0.81	0.81	0.81	80%
	Asthmatic	0.86	0.86	0.86	
CV	Control	0.80	0.79	0.80	82%
	Asthmatic	0.85	0.86	0.85	

RFE: Recursive Feature Elimination, PI: Permutation Importance, CV: coefficient of variance.



(b)

Status	Precision	Recall	F1-score	Best Accuracy
Control	0.80	0.86	0.80	86%
Asthmatic	0.89	0.89	0.86	

FIGURE 4

Performance metrics for 33 Pooled Features. Performance metrics of a trained SVM using all 33 Pooled features, showcasing (a) observed macro-averaged ROC-AUC of 0.9691 ± 0.00 and (b) precision, recall, and f1-scores for both asthmatic and control classes.

of grain classification, dosiomics in radiation therapy and x-ray imaging (De Moura et al., 2020; Placidi et al., 2021; Robert Singh and Chaudhury, 2020).

The number of samples relative to the number of features is crucial. Using a subset of eighty features for a sample size of 549 is justifiable, however, we increase the dimensionality of the dataset at the cost of computational and classification complexity. Regardless of the SVM classifiers exhibiting optimal F1, precision, and recall scores for each class when supplied with all 80 textural features, a near-perfect macro-averaged ROC of 0.98 is indicative of poor generalization. Standard approaches to address overfit models involve tasks such as eliminating outliers from the dataset, expanding the dataset size, rectifying class imbalances, and/or enhancing feature refinement. Due to sample limitations, increasing the size of the dataset is not a feasible option for our study. Using data engineering methods, such as image mirroring and stretching to grow the dataset would negatively impact the collagen morphologies unique to both the remodeled asthmatic airways and control airways. With an evenly balanced dataset of 296 images of asthmatic remodeled airways and 253 of control airways, the threat of class imbalance can be further eliminated.

To better understand the analysis of collagen morphologies using gray-level textures, it is crucial to identify how the selected features capture biological changes such as collagen fragmentation,

fiber alignment, and disease progression. As expected, all four FOS features were eliminated by every refinement method. FOS primarily represents pixel intensity distributions without spatial context and thus fails to capture the broader structural patterns essential for distinguishing healthy from asthmatic tissues.

ANOVA F-Test and RFE (our two best-performing subsets) prioritized GLCM, GLSZM, GLRLM, and NGTDM features because these metrics reflect organization and continuity (or fragmentation) of fibrillar collagen bundles. For instance, GLCM correlation measures the likelihood of repetitive patterns, which is lower in asthmatic airways where collagen fibers become disorganized and fragmented. GLSZM features (e.g., large area emphasis) relate to the size and clustering of fibrotic collagen bundles—often larger, more irregular, and fragmented in diseased states. Similarly, GLRLM features capture consecutive pixel runs along specific directions; higher short-run emphasis, for example, signifies short segments of collagen indicative of pathological breakdown.

In line with these findings, other studies using large radiomics datasets (Moura et al., 2022; Zhang et al., 2023) confirm the necessity of feature refinement when dealing with high-dimensional feature spaces. While LASSO regression is common in radiomics, we compared four different refinement methods (ANOVA, RFE, PI, and CoV) to ensure an unbiased understanding of feature relevance

in differentiating remodeling patterns in asthma. RFE's reliance on supervised models can risk overfitting, but the final chosen features still yielded robust classification performance. Conversely, filter methods like ANOVA directly measure the separation between classes, selecting features that reflect morphological distinctions (e.g., fibril alignment vs. disorganized deposition) without depending on a predictive model.

Our key objective is to explain the observed collagen remodeling rather than merely optimize classification metrics. Collagen I and III, which provide tensile strength and regulate cellular responses, often exhibit fragmented or over-deposited patterns in asthmatic airways (Burgstaller et al., 2017). Pathologic changes in collagen fibril formation can negatively impair cell polarity and alignment (Rozario and DeSimone, 2010). These changes are strongly associated with the pathogenesis and progression of airway disease (Liu Y. et al., 2021). Hence, features capturing high entropy, large fragmented zones, or reduced fiber orientation (e.g., run entropy, size zone variance, and busyness) are biologically meaningful indicators of disease progression. We found that GLSZM and GLRLM metrics were particularly sensitive to these changes, highlighting the thicker, more variably oriented bundles commonly observed in remodeled airways.

Though deep-learning methods and transfer learning have proven to be successful at denoting textural variation and outperforming classical models in accuracy, they often obscure which specific image features drive predictions. With them aim of improving classification accuracy, Davydko et al., propose the aggregation of all five textural matrices using a Feature-Constructor neural network structure comprised of an encoder, encoded features aggregator and neural networks (Davydko et al., 2021). The hybrid classifiers based on CNN, LSOE, GMDH are applied to characterize pneumonic lesion types of COVID-19. Though a total accuracy of 0.96 was achieved, the optimal aggregation of the five textural matrices remains blinded to the user, shielding the contribution of each specified feature. The authors state that though the application of all five matrices are powerful textural descriptors, they do not all necessarily apply in the context of their classification problem. Hence, insight into the exact feature subsets leveraged for high classification accuracy is necessitated (Davydko et al., 2021). In contrast, our approach uses hand-selected gray-level features to quantitatively link machine-learning outputs to actual collagen changes in the tissue microenvironment. This transparency helps validate the findings from a biological standpoint, providing a clearer picture of how collagen remodeling manifests in the ECM and influences disease severity.

To conclude, feature refinement was imperative to reduce our initial set of 80 features to a manageable subset that best reflects collagen's structural and morphological alterations. GLSZM and GLRLM in particular provided strong discriminative power, offering insight into bundle size, fragmentation, and orientation. While each texture matrix has limitations—such as GLSZM's lack of orientation or GLCM's reliance on only four directional angles—combining multiple matrices yields a robust, multifaceted assessment of fibril arrangement. Ultimately, this approach enables a clearer biological interpretation of collagen remodeling in asthma and underscores the necessity of selecting textural features that specifically capture pathological fiber changes.

Despite the novelty of employing a substantial number of gray texture features, incorporating local binary pattern textures or Law's textures could aid in addressing pertinent information gaps. Furthermore, the features computed in this study lacked scale invariance. Hence, alterations in image magnification, variations in ROI and observation window sizes, pose a threat to the calculated features. To ensure consistency, a set of images should be captured at the same magnification, as the impact of differing ROIs or observation window sizes was evident. The deliberate selection of ROIs, through either manual intervention or automated methods, could curb this impact.

All presented models (trained using all 80 features, subsets of refined features and pooled features) encountered challenges when classifying the collagen morphology of control airways. In the applied dataset, significant overlaps in the architecture of collagen in both remodeled asthmatic and control airways were evident. Even with the supply of high predictive features, the model presented challenges to separate amongst disease groups. This demonstrates the difficulty in the exploration of the cellular microenvironment when samples are not from tightly controlled sources. Biological samples display a wider range of differences in collagen network morphology within groups. The probability of control collagen being misclassified as asthmatic airways was more prevalent than the reverse case. Though it is outside of the scope of this study, it would be pertinent to explore the relation of collagen distribution amongst individuals of varying sexes, ages and differing comorbidities. The lack of standardization of the collagen morphologies amongst our control donors could lead to added complexity in the classification task. Ideally, individualistic traits such as age and sex should be accounted for as they are potential factors driving differences in the collagen microstructure. However, the current work was limited in its sample size, curbing the exploration of the aforementioned. With an expanded dataset, these characteristics will be considered in future analysis. To better support the generalization of our proposed pipeline, the employed set of textural features should be tested on an independent secondary dataset of relevant SHG images. Though it is not presented in this work, the proposed pipeline has been applied to characterize collagen-characteristics in varying SHG datasets, analyzing the role of collagen morphology and structure in age-related changes of tendons, burn effects in skin, development of atherosclerosis and infarcted hearts.

The work presented facilitates the extrapolation in assessing the effectiveness of drug treatments in reversing the fibrotic response observed during the onset of asthma. With this proof of concept, clinicians could investigate the impact of these treatments on ECM remodeling and ascertain their efficacy in redirecting collagen accumulation in presuming a non-fibrotic environment.

5 Conclusion

Extracting textural features associated with collagen remodeling in biomedical tissues offers a more comprehensive and objective analysis of fibrotic responses, especially those related to changes in airway structures. Such metrics can be used to determine the extent of airway remodeling in obstructive lung diseases such as asthma or chronic obstructive pulmonary disease and/or the effects of potential novel therapies. The combination of label-free microscopy

imaging and supervised machine learning approaches allowed us to demonstrate for the first time that gray-level-based textural variation of the fibrillar collagen can be a promising approach to identify subtle structural changes in asthmatics compared to non-asthmatics airways.

Data availability statement

The datasets presented in this article are not readily available due to the novelty of the SHG-imaged lung collagen dataset and its use for supporting research projects, the raw data will not be shared as of yet. Requests to access the datasets should be directed to Dr. Leila B. Mostaço-Guidolin, leila@sce.carleton.ca.

Ethics statement

This study was approved by the Providence Health Care Research Ethics Board (H13-02173). The studies were conducted in accordance with the local legislation and institutional requirements. The human lung tissue samples used in this study were acquired as part of a previous study for which ethical approval was obtained. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

NNK: Conceptualization, Formal Analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing—original draft, Writing—review and editing. JJAP: Conceptualization, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing—review and editing. JL: Formal Analysis, Visualization, Writing—review and editing. T-LH: Conceptualization, Data curation, Investigation, Methodology, Project administration, Resources, Validation, Writing—review and editing. RT: Visualization, Writing—review and editing, Validation. LBM-G: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing—original draft, Writing—review and editing.

References

- Altmann, A., Toloşi, L., Sander, O., and Lengauer, T. (2010). Permutation importance: a corrected feature importance measure. *Bioinformatics* 26 (10), 1340–1347. doi:10.1093/bioinformatics/btq134
- Alturkistani, H. A., Tashkandi, F. M., and Mohammadsaleh, Z. M. (2015). Histological stains: A literature review and case study. *Glob. J. Health Sci.* 8 (3), 72. doi:10.5539/gjhs.v8n3p72
- Alvarenga, A. V., Infantosi, A. F. C., Pereira, W. C. A., and Azevedo, C. M. (2010). Assessing the performance of morphological parameters in distinguishing breast tumors on ultrasound images. *Med. Eng. and Phys.* 32 (1), 49–56. doi:10.1016/j.medengphys.2009.10.007
- Amadasun, M., and King, R. (1989). Textural features corresponding to textural properties. *IEEE Trans. Syst. Man, Cybern.* 19 (5), 1264–1274. doi:10.1109/21.44046

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the Natural Sciences and Engineering Research Council (NSERC), Discovery Grant of Canada [grant number RGPIN-2021-04185] and Carleton University (Research Development Grant). NNK was supported by Asthma Canada, the Canadian Allergy, Asthma and Immunology Foundation (CAAIF) [grant number 02357-000] and the Canadian Institutes of Health Research - Institute of Circulatory and Respiratory Health (CIHR-ICRH) [grant number 202311FBD]. LBM-G and T-LH are supported by funding from the Canada Research Chairs (CRC) Program.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fbinf.2025.1539936/full#supplementary-material>

- Bergeron, C., Tulic, M. K., and Hamid, Q. (2010). Airway remodelling in asthma: from benchside to clinical practice. *Can. Respir. J.* 17 (4), e85–e93. doi:10.1155/2010/318029
- Boulet, L.-P., Laviolette, M., Turcotte, H., Cartier, A., Dugas, M., Malo, J.-L., et al. (1997). Bronchial subepithelial fibrosis correlates with airway responsiveness to methacholine. *Chest* 112 (1), 45–52. doi:10.1378/chest.112.1.45
- Bredfeldt, J. S., Liu, Y., Pehlke, C. A., Conklin, M. W., Szulcowski, J. M., Inman, D. R., et al. (2014). Computational segmentation of collagen fibers from second-harmonic generation images of breast cancer. *J. Biomed. Opt.* 19 (1), 016007. doi:10.1117/1.JBO.19.1.016007
- Buehler, M. J. (2006). Nature designs tough collagen: Explaining the nanostructure of collagen fibrils. *Proc. Natl. Acad. Sci.* 103 (33), 12285–12290. doi:10.1073/pnas.0603216103
- Burgstaller, G., Oehrle, B., Gerckens, M., White, E. S., Schiller, H. B., and Eickelberg, O. (2017). The instructive extracellular matrix of the lung: Basic composition and alterations in chronic lung disease. *Eur. Respir. J.* 50 (1), 1601805. doi:10.1183/13993003.01805-2016
- Carroll, N. G., Perry, S., Karkhanis, A., Harji, S., Butt, J., James, A. L., et al. (2000). The airway longitudinal elastic fiber network and mucosal folding in patients with asthma. *Am. J. Respir. Crit. Care Med.* 161 (1), 244–248. doi:10.1164/ajrccm.161.1.9805005
- Chabat, F., Yang, G.-Z., and Hansell, D. M. (2003). Obstructive lung diseases: Texture classification for differentiation at CT. *Radiology* 228 (3), 871–877. doi:10.1148/radiol.2283020505
- Chen, X., Nadiarynkh, O., Plotnikov, S., and Campagnola, P. J. (2012). Second harmonic generation microscopy for quantitative analysis of collagen fibrillar structure. *Nat. Protoc.* 7 (4), 654–669. doi:10.1038/nprot.2012.009
- Cicchi, R., Vogler, N., Kapsokalyvas, D., Dietzek, B., Popp, J., and Pavone, F. S. (2013). From molecular structure to tissue architecture: Collagen organization probed by SHG microscopy. *J. Biophot.* 6 (2), 129–142. doi:10.1002/jbio.201200092
- Cox, G., Kable, E., Jones, A., Fraser, I., Manconi, F., and Gorrell, M. D. (2003). 3-Dimensional imaging of collagen using second harmonic generation. *J. Struct. Biol.* 141 (1), 53–62. doi:10.1016/S1047-8477(02)00576-2
- Davydko, O., Hladkyi, Y., Linnik, M., Nosovets, O., Pavlov, V., and Nastenkov, I. (2021). “Hybrid classifiers based on CNN, LSOF, GMDH in COVID-19 pneumonic lesions types classification task,” in *2021 IEEE 16th international conference on computer Sciences and information Technologies (CSIT)*, 380–384. doi:10.1109/CSIT52700.2021.9648752
- Deka, G., Wu, W.-W., and Kao, F.-J. (2012). *In vivo* wound healing diagnosis with second harmonic and fluorescence lifetime imaging. *J. Biomed. Opt.* 18 (6), 061222. doi:10.1117/1.JBO.18.6.061222
- De Moura, L. V., Dartora, C. M., De Oliveira, C. M., Barros, R. C., and Da Silva, A. M. M. (2020). A novel approach to differentiate COVID-19 pneumonia in chest X-ray. *IEEE 20th Int. Conf. Bioinforma. Bioeng. (BIBE)*, 446–451. doi:10.1109/BIBE50027.2020.00078
- Dissanayake, K., and Md Johar, M. G. (2021). Comparative study on heart disease prediction using feature selection techniques on classification algorithms. *Appl. Comput. Intell. Soft Comput.* 2021, 1–17. doi:10.1155/2021/5581806
- Dunnill, M. S. (1960). The pathology of asthma, with special reference to changes in the bronchial mucosa. *J. Clin. Pathology* 13 (1), 27–33. doi:10.1136/jcp.13.1.27
- Elias, J. A., Zhu, Z., Chupp, G., and Homer, R. J. (1999). Airway remodeling in asthma. *J. Clin. Investigation* 104 (8), 1001–1006. doi:10.1172/JCI8124
- Engler, A. J., Humbert, P. O., Wehrle-Haller, B., and Weaver, V. M. (2009). Multiscale modeling of form and function. *Science* 324 (5924), 208–212. doi:10.1126/science.1170107
- Galloway, M. M. (1975). Texture analysis using gray level run lengths. *Comput. Graph. Image Process.* 4 (2), 172–179. doi:10.1016/S0146-664X(75)80008-6
- Gibbs, P., and Turnbull, L. W. (2003). Textural analysis of contrast-enhanced MR images of the breast. *Magnetic Reson. Med.* 50 (1), 92–98. doi:10.1002/mrm.10496
- Grothausmann, R., Knudsen, L., Ochs, M., and Mühlfeld, C. (2017). Digital 3D reconstructions using histological serial sections of lung tissue including the alveolar capillary network. *Am. J. Physiology-Lung Cell. Mol. Physiology* 312 (2), L243–L257. doi:10.1152/ajplung.00326.2016
- Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). No title found. *Mach. Learn.* 46 (1/3), 389–422. doi:10.1023/A:1012487302797
- Han, M., Giese, G., and Bille, J. F. (2005). Second harmonic generation imaging of collagen fibrils in cornea and sclera. *Opt. Express* 13 (15), 5791. doi:10.1364/OPEX.13.005791
- Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). Textural features for image classification. *IEEE Trans. Syst. Man, Cybern.* SMC-3 (6), 610–621. doi:10.1109/TSMC.1973.4309314
- Herlidou, S., Grebe, R., Grados, F., Leuyer, N., Fardellone, P., and Meyer, M.-E. (2004). Influence of age and osteoporosis on calcaneus trabecular bone structure: A preliminary *in vivo* MRI study by quantitative texture analysis. *Magn. Reson. Imaging* 22 (2), 237–243. doi:10.1016/j.mri.2003.07.007
- Hogg, J. C. (1997). The pathology of asthma. *APMIS* 105 (7–12), 735–745. doi:10.1111/j.1699-0463.1997.tb05079.x
- Hogg, J. C., Chu, F., Utokaparch, S., Woods, R., Elliott, W. M., Buzatu, L., et al. (2004). The nature of small-airway obstruction in chronic obstructive pulmonary disease. *N. Engl. J. Med.* 350 (26), 2645–2653. doi:10.1056/NEJMoa032158
- Holgate, S. T. (2008). Pathogenesis of asthma. *Clin. and Exp. Allergy* 38 (6), 872–897. doi:10.1111/j.1365-2222.2008.02971.x
- Hoshino, M., Nakamura, Y., Sim, J., Shimojo, J., and Isogai, S. (1998). Bronchial subepithelial fibrosis and expression of matrix metalloproteinase-9 in asthmatic airway inflammation. *J. Allergy Clin. Immunol.* 102(5), 783–788. doi:10.1016/S0091-6749(98)70018-1
- Hu, W., Li, H., Wang, C., Gou, S., and Fu, L. (2012). Characterization of collagen fibers by means of texture analysis of second harmonic generation images using orientation-dependent gray level co-occurrence matrix method. *J. Biomed. Opt.* 17 (2), 026007. doi:10.1117/1.JBO.17.2.026007
- Huber, H. L. (1922). THE PATHOLOGY OF BRONCHIAL ASTHMA. *Archives Intern. Med.* 30 (6), 689. doi:10.1001/archinte.1922.00110120002001
- Israel, J. S., Esquibel, C. R., Dingle, A. M., Liu, Y., Keikhosravi, A., Pisaniello, J. A., et al. (2017). Quantification of collagen organization after nerve repair. *Plastic Reconstr. Surg. - Glob. Open* 5 (12), e1586. doi:10.1097/GOX.0000000000001586
- James, A. L., Maxwell, P. S., Pearce-Pinto, G., Elliot, J. G., and Carroll, N. G. (2002). The relationship of reticular basement membrane thickness to airway wall remodeling in asthma. *Am. J. Respir. Crit. Care Med.* 166 (12), 1590–1595. doi:10.1164/rccm.2108069
- Javeed, A., Qamar, S., Ali, S., Mustafa, M. A. T., Nusrat, A., and Ghauri, S. K. (2021). Histological stains in the past, present, and future. *Cureus* 13, e18486. doi:10.7759/cureus.18486
- Keikhosravi, A., Li, B., Liu, Y., and Eliceiri, K. W. (2020). Intensity-based registration of bright-field and second-harmonic generation images of histopathology tissue sections. *Biomed. Opt. Express* 11 (1), 160. doi:10.1364/BOE.11.000160
- Ko, A. C. T., Ridsdale, A., Pegoraro, A. F., Smith, M. S. D., Mostaço-Guidolin, L. B., Hewko, M. D., et al. (2009). Label-free imaging of arterial tissues using photonic crystal fiber (PCF) based nonlinear optical microscopic system. 718204. doi:10.1117/12.809361
- Ko, A. C. T., Ridsdale, A., Smith, M. S. D., Mostaço-Guidolin, L. B., Hewko, M. D., Pegoraro, A. F., et al. (2010). Multimodal nonlinear optical imaging of atherosclerotic plaque development in myocardial infarction-prone rabbits. *J. Biomed. Opt.* 15 (2), 020501. doi:10.1117/1.3353960
- Kunchur, N. N., and Mostaço-Guidolin, L. B. (2022). Development of an image classification pipeline for atherosclerotic plaques assessment using supervised machine learning. *BMC Bioinforma.* 23 (1), 542. doi:10.1186/s12859-022-05059-1
- Liu, L., Stephens, B., Bergman, M., May, A., and Chiang, T. (2021). Role of collagen in airway mechanics. *Bioengineering* 8 (1), 13. doi:10.3390/bioengineering8010013
- Liu, Y., Dong, Y., Si, L., Meng, R., Dong, Y., and Ma, H. (2021). Comparison between image texture and polarization features in histopathology. *Biomed. Opt. Express* 12 (3), 1593. doi:10.1364/BOE.416382
- Maronna, R. A., Martin, R. D., and Yohai, V. J. (2006). *Robust statistics: theory and methods*. 1st ed. Wiley. doi:10.1002/0470010940
- Martin, T. P., Norris, G., McConnell, G., and Currie, S. (2013). A novel approach for assessing cardiac fibrosis using label-free second harmonic generation. *Int. J. Cardiovasc. Imaging* 29 (8), 1733–1740. doi:10.1007/s10554-013-0270-2
- Monaghan, M. G., Kroll, S., Brucker, S. Y., and Schenke-Layland, K. (2016). Enabling multiphoton and second harmonic generation imaging in paraffin-embedded and histologically stained sections. *Tissue Eng. Part C. Methods* 22 (6), 517–523. doi:10.1089/ten.tec.2016.0071
- Mostaço-Guidolin, L. B., Osei, E. T., Ullah, J., Hajimohammadi, S., Fouadi, M., Li, X., et al. (2019). Defective fibrillar collagen organization by fibroblasts contributes to airway remodeling in asthma. *Am. J. Respir. Crit. Care Med.* 200 (4), 431–443. doi:10.1164/rccm.201810-1855OC
- Mostaço-Guidolin, L. B., Yang, C. X., and Hackett, T.-L. (2021). Pulmonary vascular remodeling is an early feature of fatal and nonfatal asthma. *Am. J. Respir. Cell Mol. Biol.* 65 (1), 114–118. doi:10.1165/rcmb.2020-0339LE
- Moura, L. V. D., Mattjie, C., Dartora, C. M., Barros, R. C., and Marques Da Silva, A. M. (2022). Explainable machine learning for COVID-19 pneumonia classification with texture-based features extraction in chest radiography. *Front. Digital Health* 3, 662343. doi:10.3389/fdgh.2021.662343
- Placidi, L., Gioscio, E., Garibaldi, C., Rancati, T., Fanizzi, A., Maestri, D., et al. (2021). A multicenter evaluation of dosiomics features reproducibility, stability and sensitivity. *Cancers* 13 (15), 3835. doi:10.3390/cancers13153835
- Pologruto, T. A., Sabatini, B. L., and Svoboda, K. (2003). ScanImage: flexible software for operating laser scanning microscopes. *Biomed. Eng. OnLine* 2 (1), 13. doi:10.1186/1475-925X-2-13
- Poole, J. J. A., and Mostaço-Guidolin, L. B. (2021). Optical microscopy and the extracellular matrix structure: a review. *Cells* 10 (7), 1760. doi:10.3390/cells10071760

- Remeseiro, B., and Bolon-Canedo, V. (2019). A review of feature selection methods in medical applications. *Comput. Biol. Med.* 112, 103375. doi:10.1016/j.compbiomed.2019.103375
- Robert Singh, K., and Chaudhury, S. (2020). Comparative analysis of texture feature extraction techniques for rice grain classification. *IET Image Process.* 14 (11), 2532–2540. doi:10.1049/iet-ipr.2019.1055
- Rozario, T., and DeSimone, D. W. (2010). The extracellular matrix in development and morphogenesis: A dynamic view. *Dev. Biol.* 341 (1), 126–140. doi:10.1016/j.ydbio.2009.10.026
- Sabo, A. R., Winfree, S., Bledsoe, S. B., Phillips, C. L., Lingeman, J. E., Eadon, M. T., et al. (2022). Label-free imaging of non-deparaffinized sections of the human kidney to determine tissue quality and signatures of disease. *Physiol. Rep.* 10 (3), e15167. doi:10.14814/phy2.15167
- Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9 (7), 671–675. doi:10.1038/nmeth.2089
- scikit-learn, developers BSD L (2007a). F_classif. F_classif. Available online at: https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.f_classif.html.
- scikit-learn, developers BSD L (2007b). Permutation_importance. Permutation_importance. Available online at: https://scikit-learn.org/stable/modules/generated/sklearn.inspection.permutation_importance.html.
- scikit-learn, developers BSD L (2007c). SVC. SVC. Available online at: <https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html>.
- scikit-learn developers, BSD License (2007). RFE. RFE. Available online at: https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.RFE.html.
- Song, J., Zhang, Y., Cheng, J., Wang, S., Liu, Z., and Sun, D. (2022). Non-invasive quantitative diagnosis of liver fibrosis with an artificial neural network. *Neural Comput. Appl.* 34 (9), 6733–6744. doi:10.1007/s00521-021-06257-3
- Sun, C., and Wee, W. G. (1983). Neighboring gray level dependence matrix for texture classification. *Comput. Vis. Graph. Image Process.* 23 (3), 341–352. doi:10.1016/0734-189X(83)90032-4
- Sun, W., Chang, S., Tai, D. C. S., Tan, N., Xiao, G., Tang, H., et al. (2008). Nonlinear optical microscopy: Use of second harmonic generation and two-photon microscopy for automated quantitative liver fibrosis studies. *J. Biomed. Opt.* 13 (6), 064010. doi:10.1117/1.3041159
- Thibault, G., Angulo, J., and Meyer, F. (2014). Advanced statistical matrices for texture characterization: Application to cell classification. *IEEE Trans. Biomed. Eng.* 61 (3), 630–637. doi:10.1109/TBME.2013.2284600
- Tilbury, K., Hocker, J., Wen, B. L., Sandbo, N., Singh, V., and Campagnola, P. J. (2014). Second harmonic generation microscopy analysis of extracellular matrix changes in human idiopathic pulmonary fibrosis. *J. Biomed. Opt.* 19 (8), 086014. doi:10.1117/1.JBO.19.8.086014
- Vijayaraghavan, S., Huq, R., and Hausman, M. R. (2014). Methods of peripheral nerve tissue preparation for second harmonic generation imaging of collagen fibers. *Methods* 66 (2), 246–255. doi:10.1016/j.ymeth.2013.08.012
- Wu, S., Peng, Y., Hu, L., Zhang, X., and Li, H. (2016). Classification and recognition of texture collagen obtaining by multiphoton microscope with neural network analysis. *J. Phys. Conf. Ser.* 680, 012014. doi:10.1088/1742-6596/680/1/012014
- Zeiger, R. S., Dawsonb, C., and Weiss, S. (1999). Relationships between duration of asthma and asthma severity among children in the Childhood Asthma Management Program (CAMP). *J. Allergy Clin. Immunol.* 103 (3), 376–386. doi:10.1016/S0091-6749(99)70460-4
- Zhang, Y.-P., Zhang, X.-Y., Cheng, Y.-T., Li, B., Teng, X.-Z., Zhang, J., et al. (2023). Artificial intelligence-driven radiomics study in cancer: The role of feature engineering and modeling. *Mil. Med. Res.* 10 (1), 22. doi:10.1186/s40779-023-00458-8